

10 Rue de la Paix 25 JUN 2004

特許協力条約

PCT

国際予備審査報告

REC'D 26 SEP 2003

WIPO

PCT

(法第12条、法施行規則第56条)
(PCT36条及びPCT規則70)

出願人又は代理人 の書類記号 PCLA-14604	今後の手続きについては、国際予備審査報告の送付通知（様式PCT/IPEA/416）を参照すること。		
国際出願番号 PCT/JP03/00011	国際出願日 (日.月.年) 06.01.03	優先日 (日.月.年) 28.12.01	
国際特許分類 (IPC) Int. C17 G06F17/30, C12Q1/68, C12N15/00, G01N33/50			
出願人（氏名又は名称） セレスター・レキシコ・サイエンシズ株式会社			

1. 国際予備審査機関が作成したこの国際予備審査報告を法施行規則第57条（PCT36条）の規定に従い送付する。

2. この国際予備審査報告は、この表紙を含めて全部で 4 ページからなる。

この国際予備審査報告には、附属書類、つまり補正されて、この報告の基礎とされた及び／又はこの国際予備審査機関に対して訂正を含む明細書、請求の範囲及び／又は図面も添付されている。
(PCT規則70.16及びPCT実施細則第607号参照)
この附属書類は、全部で 24 ページである。

3. この国際予備審査報告は、次の内容を含む。

- I 国際予備審査報告の基礎
- II 優先権
- III 新規性、進歩性又は産業上の利用可能性についての国際予備審査報告の不作成
- IV 発明の単一性の欠如
- V PCT35条(2)に規定する新規性、進歩性又は産業上の利用可能性についての見解、それを裏付けるための文献及び説明
- VI ある種の引用文献
- VII 国際出願の不備
- VIII 国際出願に対する意見

国際予備審査の請求書を受理した日 22.05.03	国際予備審査報告を作成した日 09.09.03
名称及びあて先 日本国特許庁 (IPEA/JP) 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号	特許庁審査官（権限のある職員） 高瀬 勤 (印)
	5M 9069
	電話番号 03-3581-1101 内線 3597

I. 国際予備審査報告の基礎

1. この国際予備審査報告は下記の出願書類に基づいて作成された。(法第6条(PCT14条)の規定に基づく命令に応答するために提出された差し替え用紙は、この報告書において「出願時」とし、本報告書には添付しない。PCT規則70.16, 70.17)

出願時の国際出願書類

明細書 第 1-4, 7, 8, 11, 12, 16-50, 53 ページ、
明細書 第 _____ ページ、
明細書 第 5-6/1, 9-10/1, 13-15/1 51-52/1 ページ、

出願時に提出されたもの
国際予備審査の請求書と共に提出されたもの
01.09.03 付の書簡と共に提出されたもの

請求の範囲 第 3-6, 9-, 12, 15-18, 21-24 項、
請求の範囲 第 _____ 項、
請求の範囲 第 _____ 項、
請求の範囲 第 1, 2, 7, 8, 13, 14, 19, 20 項、

出願時に提出されたもの
PCT19条の規定に基づき補正されたもの
国際予備審査の請求書と共に提出されたもの
01.09.03 付の書簡と共に提出されたもの

図面 第 1-23 ページ/図、
図面 第 _____ ページ/図、
図面 第 _____ ページ/図、

出願時に提出されたもの
国際予備審査の請求書と共に提出されたもの
付の書簡と共に提出されたもの

明細書の配列表の部分 第 _____ ページ、
明細書の配列表の部分 第 _____ ページ、
明細書の配列表の部分 第 _____ ページ、

出願時に提出されたもの
国際予備審査の請求書と共に提出されたもの
付の書簡と共に提出されたもの

2. 上記の出願書類の言語は、下記に示す場合を除くほか、この国際出願の言語である。

上記の書類は、下記の言語である _____ 語である。

国際調査のために提出されたPCT規則23.1(b)にいう翻訳文の言語
 PCT規則48.3(b)にいう国際公開の言語
 国際予備審査のために提出されたPCT規則55.2または55.3にいう翻訳文の言語

3. この国際出願は、ヌクレオチド又はアミノ酸配列を含んでおり、次の配列表に基づき国際予備審査報告を行った。

この国際出願に含まれる書面による配列表
 この国際出願と共に提出された磁気ディスクによる配列表
 出願後に、この国際予備審査（または調査）機関に提出された書面による配列表
 出願後に、この国際予備審査（または調査）機関に提出された磁気ディスクによる配列表
 出願後に提出した書面による配列表が出願時における国際出願の開示の範囲を超える事項を含まない旨の陳述書の提出があった
 書面による配列表に記載した配列と磁気ディスクによる配列表に記録した配列が同一である旨の陳述書の提出があった。

4. 振正により、下記の書類が削除された。

明細書 第 _____ ページ
 請求の範囲 第 _____ 項
 図面 図面の第 _____ ページ/図

5. この国際予備審査報告は、補充欄に示したように、補正が出願時における開示の範囲を越えてされたものと認められるので、その補正がされなかったものとして作成した。(PCT規則70.2(c) この補正を含む差し替え用紙は上記1.における判断の際に考慮しなければならず、本報告に添付する。)

V. 新規性、進歩性又は産業上の利用可能性についての法第12条 (PCT35条(2)) に定める見解、それを裏付ける文献及び説明

1. 見解

新規性 (N)

請求の範囲 1-24

有

請求の範囲

無

進歩性 (IS)

請求の範囲 3, 6, 9, 12, 15, 18, 21, 24

有

請求の範囲 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19, 20, 22, 23

無

産業上の利用可能性 (IA)

請求の範囲 1-24

有

請求の範囲

無

2. 文献及び説明 (PCT規則70.7)

国際調査報告で引用した

文献1：吉沢 他, RNA二次構造予測支援システムの開発と実装, 人工知能学会人工知能基礎論研究会(第32回)資料, 1998.03.26, p. 29-35, 特に, p. 31

文献2：UEMURA, Y. et al. Grammatically Modeling and Predicting RNA Secondary Structures. Genome Informatics Series, Proceedings Genome Informatics Workshop, 1995.12.11, No. 6, p. 67-76

請求の範囲1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19, 20, 22, 23は文献1より進歩性を有しない。

説明：

文献1 (p. 31) には、RNA入力配列を記述することが可能な構文解析木を求め、自由エネルギー値の小さなものから上位いくつかをRNAの二次構造の候補として出力することが記載されている。

請求の範囲1, 7, 13, 19に関して

文献1記載のものも、予測結果をトポジカルな構造情報による分類を行うことが有効であることも記載されている (p. 34右欄第12-14行)。

請求の範囲2, 8, 14, 20に関して

与えられた二次構造を取り得るRNA配列を検索することは当業者が適宜なし得る事項にすぎない。

請求の範囲4, 10, 16, 22に関して

cDNA配列とRNA配列は対応するものであるから、DNA配列入力によりRNA配列変換してから、文献1に記載された二次構造予測を行うことは当業者にとって容易である。

請求項の範囲5, 11, 17, 23に関して

二次構造の類似していれば、RNAも類似していると予測できるから、二次構造の適合度をRNA配列間の類似度とすることは当業者が適宜なし得る事項にすぎない。

補充欄（いずれかの欄の大きさが足りない場合に使用すること）

第 V 欄の続き

請求の範囲3, 6, 9, 12, 15, 18, 21, 24は文献1, 2に対して新規性、進歩性を有する。

説明：

構造トポロジーとRNA配列とを二次元マトリックスで表示し、格子部分をマークすることは文献1, 2には記載も示唆もされていない。

発明の開示

本発明にかかるRNA配列解析装置は、RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納する文法格納手段と、RNA配列を複数の上記生成文法に適用して構文解析木を導出する構文解析手段と、上記構文解析手段にて導出された上記構文解析木に対して適合度の計算を行う適合度計算手段と、上記適合度計算手段により計算された上記適合度のうち予め定めた条件を満たす適合度である上記構文解析木を上記適合度が高い順にソートするソート手段と、上記ソート手段によりソートされた上記構文解析木を上記RNA配列の二次構造の候補として出力する出力手段とを備えたことを特徴とする。

この装置によれば、RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納し、RNA配列を複数の生成文法に適用して構文解析木を導出し、導出された構文解析木に対して適合度の計算し、計算された適合度のうち予め定めた条件を満たす適合度である構文解析木を適合度が高い順にソートし、ソートされた構文解析木をRNA配列の二次構造の候補として出力するので、一配列に対して多文法で構文解析を行うことができるようになる。すなわち、各生成文法に対してそれぞれ構文解析し適合度計算を行い適合度を得る。その結果、生成文法ごとに適合度が得られることになり、それらの適合度をソートすることによって生成文法に順位を付ける。これにより、生成文法に対する構造トポロジーにも順位が付けられることになるので、最終的にRNA配列が取り得る可能性の高い順に構造トポロジーを確認することができるようになる。

つぎの発明にかかるRNA配列解析装置は、RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納する文法格納手段と、複数のRNA配列を予め定めた上記生成文法に適用して構文解析木を導出する構文解析手段と、上記構文解析手段にて導出された上記構文解析木に対して適合度の計算を行う適合度計算手段と、上記適合度計算手段により計算された上

記適合度のうち予め定めた条件を満たす適合度である上記構文解析木を導出した上記RNA配列を上記予め定めた上記生成文法が適合する上記構造トポロジーの二次構造を持つRNA配列の候補として出力する出力手段とを備えたことを特徴とする。

5 この装置によれば、RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納し、複数のRNA配列を予め定めた生成文法に適用して構文解析木を導出し、導出された構文解析木に対して適合度の計算を行い、計算された適合度のうち予め定めた条件を満たす適合度である構文解析木を導出したRNA配列を予め定めた生成文法が適合する上記構造トポロジーの二次構造を持つRNA配列の候補として出力する上記構造トポロジーの二次構造を持つRNA配列の候補として出力するので、多配列に対して一文法で構文解析を行うことができるようになる。すなわち、与えられた特定の構造トポロジーに対し、対応する生成文法を取得し、これを用いてRNA配列データベースに格納されているすべてまたは一部のRNA配列をそれぞれ構文解析し、ある閾値以下の適合度で構文解析に成功したRNA配列群を結果として出力する。

10 これにより、与えられた特定の二次構造を取り得るようなRNA配列を検索することができるようになる。

つぎの発明にかかるRNA配列解析装置は、RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納する文法格納手段と、RNA配列を上記生成文法に適用して構文解析木を導出する構文解析手段と、上記構文解析手段にて導出された上記構文解析木に対して適合度の計算を行う適合度計算手段と、上記適合度計算手段により計算された上記適合度のうち予め定めた条件を満たす適合度である上記構文解析木を導出した上記RNA配列を抽出する抽出手段と、上記構造トポロジーと上記RNA配列とを二次元マトリックスで表示し、上記二次元マトリックスにおいて上記抽出手段にて抽出された上記RNA配列と上記構造トポロジーに対応する格子部分をマークすることにより、上記RNA配列間で共通に有する構造トポロジーを可視化する共通構造マトリックス作成手段とを備えたことを特徴とする。

この装置によれば、RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納し、RNA配列を生成文法に適用して構文解析木を導出し、導出された構文解析木に対して適合度の計算を行い、計算された適合度のうち予め定めた条件を満たす適合度である構文解析木を導出したR

文解析木を導出し、導出された構文解析木に対して適合度の計算を行い、計算された適合度のうち予め定めた条件を満たす適合度である構文解析木を導出したRNA配列を抽出し、構造トポロジーとRNA配列とを二次元マトリックスで表示し、二次元マトリックスにおいて抽出されたRNA配列と構造トポロジーに対応する格子部分に適合度を表示する適合度マトリックスを作成し、適合度マトリックスについて、適合度により構造トポロジーをソートし、他のRNA配列について当該ソートされた構造トポロジーの順番に対応する生成文法により構文解析を行い適合度が最大となる構文解析木を求め、予め定めた条件を満たす適合度を持つ構文解析木に対応する他のRNA配列を抽出するので、共通の構造を持つRNA配列を容易に発見することができるようになる。

また、本発明はRNA配列解析方法に関するものであり、本発明にかかるRNA配列解析方法は、RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納する文法格納ステップと、RNA配列を複数の上記生成文法に適用して構文解析木を導出する構文解析ステップと、上記構文解析ステップにて導出された上記構文解析木に対して適合度の計算を行う適合度計算ステップと、上記適合度計算ステップにより計算された上記適合度のうち予め定めた条件を満たす適合度である上記構文解析木を上記適合度が高い順にソートするソートステップと、上記ソートステップによりソートされた上記構文解析木を上記RNA配列の二次構造の候補として出力する出力ステップとを含むことを特徴とする。

この方法によれば、RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納し、RNA配列を複数の生成文法に適用して構文解析木を導出し、導出された構文解析木に対して適合度を計算し、計算された適合度のうち予め定めた条件を満たす適合度である構文解析木を適合度が高い順にソートし、ソートされた構文解析木をRNA配列の二次構造の候補として出力するので、一配列に対して多文法で構文解析を行うようになる。すなわち、各生成文法に対してそれぞれ構文解析し適合度計算を行い適合度

を得る。その結果、生成文法ごとに適合度が得られることになり、それらの適合度をソートすることによって生成文法に順位を付ける。これにより、生成文法に対する構造トポロジーにも順位が付けられることになるので、最終的にRNA配列が取り得る可能性の高い順に構造トポロジーを確認することができるようになる。

5 つぎの発明にかかるRNA配列解析方法は、RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納する文法格納ステップと、複数のRNA配列を予め定めた上記生成文法に適用して構文解析木を導出する構文解析ステップと、上記構文解析ステップにて導出された上記構文解析木に対して適合度の計算を行う適合度計算ステップと、上記適合度計算ステップにより計算された上記適合度のうち予め定めた条件を満たす適合度である上記構文解析木を導出した上記RNA配列を上記予め定めた上記生成文法が適合する上記構造トポロジーの二次構造を持つRNA配列の候補として出力する出力ステップとを含むことを特徴とする。

10 15 この方法によれば、RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納し、複数のRNA配列を予め定めた生成文法に適用して構文解析木を導出し、導出された構文解析木に対して適合度の計算を行い、計算された適合度のうち予め定めた条件を満たす適合度である構文解析木を導出したRNA配列を予め定めた生成文法が適合する上記構造トポロジーの二次構造を持つRNA配列の候補として出力するので、多配列に対して一文法で構文解析を行うことができるようになる。すなわち、与えられた特定の構造トポロジーに対し、対応する生成文法を取得し、これを用いてRNA配列データベースに格納されているすべてまたは一部のRNA配列をそれぞれ構文解析し、ある閾値以下の適合度で構文解析に成功したRNA配列群を結果として出力する。

20 25 これにより、与えられた特定の二次構造を取り得るようなRNA配列を検索することができるようになる。

つぎの発明にかかるRNA配列解析方法は、RNA二次構造の構造トポロジー

10/1

と、当該構造トポロジーに適合する生成文法とを対応付けて格納する文法格納ステップと、RNA配列を上記生成文法に適用して構文解析木を導出する構文解析ステップと、上記構文解析ステップにて導出された上記構文解析木に対して適合

A配列とを二次元マトリックスで表示し、上記二次元マトリックスにおいて上記抽出ステップにて抽出された上記RNA配列と上記構造トポロジーに対応する格子部分に上記適合度を表示する適合度マトリックスを作成する適合度マトリックス作成ステップと、上記適合度マトリックス作成ステップにて作成された上記適合度マトリックスについて、上記適合度により上記構造トポロジーをソートし、他のRNA配列について当該ソートされた上記構造トポロジーの順番に対応する上記生成文法により構文解析を行い上記適合度が最大となる上記構文解析木を求め、予め定めた条件を満たす上記適合度を持つ上記構文解析木に対応する上記他のRNA配列を抽出する共通構造抽出ステップとを含むことを特徴とする。

10 この方法によれば、RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納し、RNA配列を生成文法に適用して構文解析木を導出し、導出された構文解析木に対して適合度の計算を行い、計算された適合度のうち予め定めた条件を満たす適合度である構文解析木を導出したRNA配列を抽出し、構造トポロジーとRNA配列とを二次元マトリックスで表示し、二次元マトリックスにおいて抽出されたRNA配列と構造トポロジーに対応する格子部分に適合度を表示する適合度マトリックスを作成し、適合度マトリックスについて、適合度により構造トポロジーをソートし、他のRNA配列について当該ソートされた構造トポロジーの順番に対応する生成文法により構文解析を行い適合度が最大となる構文解析木を求め、予め定めた条件を満たす適合度を持つ構文解析木に対応する他のRNA配列を抽出するので、共通の構造を持つRNA配列を容易に発見することができるようになる。

また、本発明はRNA配列解析方法をコンピュータに実行させるプログラムに関するものであり、本発明にかかるプログラムは、RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納する文法格納ステップと、RNA配列を複数の上記生成文法に適用して構文解析木を導出する構文解析ステップと、上記構文解析ステップにて導出された上記構文解析木に対して適合度の計算を行う適合度計算ステップと、上記適合度計算ステップによ

り計算された上記適合度のうち予め定めた条件を満たす適合度である上記構文解析木を上記適合度が高い順にソートするソートステップと、上記ソートステップによりソートされた上記構文解析木を上記RNA配列の二次構造の候補として出力する出力ステップとを含むことを特徴とする。

5 このプログラムによれば、RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納し、RNA配列を複数の生成文法に適用して構文解析木を導出し、導出された構文解析木に対して適合度の計算し、計算された適合度のうち予め定めた条件を満たす適合度である構文解析木を適合度が高い順にソートし、ソートされた構文解析木をRNA配列の二次構造の候補
10 として出力するので、一配列に対して多文法で構文解析を行うことができるようになる。すなわち、各生成文法に対してそれぞれ構文解析し適合度計算を行い適合度を得る。その結果、生成文法ごとに適合度が得られることになり、それらの適合度をソートすることによって生成文法に順位を付ける。これにより、生成文法に対する構造トポロジーにも順位が付けられることになるので、最終的にRNA
15 配列が取り得る可能性の高い順に構造トポロジーを確認することができるようになる。

つきの発明にかかるプログラムは、RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納する文法格納ステップと、複数のRNA配列を予め定めた上記生成文法に適用して構文解析木を導出する構文解析ステップと、上記構文解析ステップにて導出された上記構文解析木に対して適合度の計算を行う適合度計算ステップと、上記適合度計算ステップにより計算された上記適合度のうち予め定めた条件を満たす適合度である上記構文解析木を導出した上記RNA配列を上記予め定めた上記生成文法が適合する上記構造トポロジーの二次構造を持つRNA配列の候補として出力する出力ステップとを含むことを特徴とする。

このプログラムによれば、RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納し、複数のRNA配列を予め定め

た生成文法に適用して構文解析木を導出し、導出された構文解析木に対して適合度の計算を行い、計算された適合度のうち予め定めた条件を満たす適合度である構文解析木を導出したRNA配列を予め定めた生成文法が適合する上記構造トポロジーの二次構造を持つRNA配列の候補として出力するので、多配列に対して一文法で構文解析を行うことができるようになる。すなわち、与えられた特定の構造トポロジーに対し、対応する生成文法を取得し、これを用いてRNA配列データベースに格納されているすべてまたは一部のRNA配列をそれぞれ構文解析し、ある閾値以下の適合度で構文解析に成功したRNA配列群を結果として出力する。これにより、与えられた特定の二次構造を取り得るようなRNA配列を検索することができるようになる。

つぎの発明にかかるプログラムは、RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納する文法格納ステップと、RNA配列を上記生成文法に適用して構文解析木を導出する構文解析ステップと、上記構文解析ステップにて導出された上記構文解析木に対して適合度の計算を行う適合度計算ステップと、上記適合度計算ステップにより計算された上記適合度のうち予め定めた条件を満たす適合度である上記構文解析木を導出した上記RNA配列を抽出する抽出ステップと、上記構造トポロジーと上記RNA配列とを二次元マトリックスで表示し、上記二次元マトリックスにおいて上記抽出ステップにて抽出された上記RNA配列と上記構造トポロジーに対応する格子部分をマークすることにより、上記RNA配列間で共通に有する構造トポロジーを可視化する共通構造マトリックス作成ステップとを含むことを特徴とする。

このプログラムによれば、RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納し、RNA配列を生成文法に適用して構文解析木を導出し、導出された構文解析木に対して適合度の計算を行い、計算された適合度のうち予め定めた条件を満たす適合度である構文解析木を導出したRNA配列を抽出し、構造トポロジーとRNA配列とを二次元マトリックスで表示し、二次元マトリックスにおいて抽出されたRNA配列と構造トポロジー

に対応する格子部分をマークすることにより、RNA配列間で共通に有する構造トポロジーを可視化するので、RNA配列間の共通構造を容易に発見することが

成、読み取り手順、あるいは、読み取り後のインストール手順等については、周知の構成や手順を用いることができる。

また、ネットワーク300は、RNA配列解析装置100と外部システム200とを相互に接続する機能を有し、例えば、インターネットや、イントラネットや、LAN(有線／無線の双方を含む)や、VANや、パソコン通信網や、公衆電話網(アナログ／デジタルの双方を含む)や、専用回線網(アナログ／デジタルの双方を含む)や、CATV網や、IMT2000方式、GSM方式またはPDC/PDC-P方式等の携带回線交換網／携帯パケット交換網や、無線呼出網や、Bluetooth等の局所無線網や、PHS網や、CS、BSまたはISDB等の衛星通信網等のうちいずれかを含んでもよい。すなわち、本システムは、有線・無線を問わず任意のネットワークを介して、各種データを送受信することができる。

以上詳細に説明したように、本発明によれば、RNA二次構造の構造トポジーと、当該構造トポジーに適合する生成文法とを対応付けて格納し、RNA配列を複数の生成文法に適用して構文解析木を導出し、導出された構文解析木に対して適合度を計算し、計算された適合度のうち予め定めた条件を満たす適合度である構文解析木を適合度が高い順にソートし、ソートされた構文解析木をRNA配列の二次構造の候補として出力するので、一配列に対して多文法で構文解析を行うことができるようになる。すなわち、各生成文法に対してそれぞれ構文解析し適合度計算を行い適合度を得る。その結果、生成文法ごとに適合度が得られることになり、それらの適合度をソートすることによって生成文法に順位を付ける。これにより、生成文法に対する構造トポジーにも順位が付けられることになるので、最終的にRNA配列が取り得る可能性の高い順に構造トポジーを確認することができるRNA配列解析装置、RNA配列解析方法、プログラム、および、記録媒体を提供することができる。

また、本発明によれば、RNA二次構造の構造トポジーと、当該構造トポジーに適合する生成文法とを対応付けて格納し、複数のRNA配列を予め定めた

生成文法に適用して構文解析木を導出し、導出された構文解析木に対して適合度の計算を行い、計算された適合度のうち予め定めた条件を満たす適合度である構文解析木を導出したRNA配列を予め定めた生成文法が適合する上記構造トポロジーの二次構造を持つRNA配列の候補として出力するので、多配列に対して一文法で構文解析を行うことができるようになる。すなわち、与えられた特定の構造トポロジーに対し、対応する生成文法を取得し、これを用いてRNA配列データベースに格納されているすべてまたは一部のRNA配列をそれぞれ構文解析し、ある閾値以下の適合度で構文解析に成功したRNA配列群を結果として出力する。これにより、与えられた特定の二次構造を取り得るようなRNA配列を検索することができるRNA配列解析装置、RNA配列解析方法、プログラム、および、記録媒体を提供することができる。

また、本発明によれば、RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納し、RNA配列を生成文法に適用して構文解析木を導出し、導出された構文解析木に対して適合度の計算を行い、計算された適合度のうち予め定めた条件を満たす適合度である構文解析木を導出したRNA配列を抽出し、構造トポロジーとRNA配列とを二次元マトリックスで表示し、二次元マトリックスにおいて抽出されたRNA配列と構造トポロジーに対応する格子部分をマークすることにより、RNA配列間で共通に有する構造トポロジーを可視化するので、RNA配列間の共通構造を容易に発見することができるRNA配列解析装置、RNA配列解析方法、プログラム、および、記録媒体を提供することができる。

また、本発明によれば、RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納し、利用者が入力したDNA配列から転写されるRNA配列を作成し、作成されたRNA配列に対して生成文法を適用して構文解析木を導出し、導出された構文解析木に対して適合度の計算を行い、計算された適合度のうち予め定めた条件を満たす適合度である構文解析木を導出したRNA配列に対応するDNA配列部分を遺伝子の候補として予測するので、

D N A配列のうち既知のトポロジーを有する可能性のあるR N A配列に対応する

請求の範囲

1. (補正後) RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納する文法格納手段と、

5 RNA配列を複数の上記生成文法に適用して構文解析木を導出する構文解析手段と、

上記構文解析手段にて導出された上記構文解析木に対して適合度の計算を行う適合度計算手段と、

10 上記適合度計算手段により計算された上記適合度のうち予め定めた条件を満たす適合度である上記構文解析木を上記適合度が高い順にソートするソート手段と、

上記ソート手段によりソートされた上記構文解析木を上記RNA配列の二次構造の候補として出力する出力手段と、

を備えたことを特徴とするRNA配列解析装置。

15 2. (補正後) RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納する文法格納手段と、

複数のRNA配列を予め定めた上記生成文法に適用して構文解析木を導出する構文解析手段と、

20 上記構文解析手段にて導出された上記構文解析木に対して適合度の計算を行う適合度計算手段と、

上記適合度計算手段により計算された上記適合度のうち予め定めた条件を満たす適合度である上記構文解析木を導出した上記RNA配列を上記予め定めた上記生成文法が適合する上記構造トポロジーの二次構造を持つRNA配列の候補として出力する出力手段と、

25 を備えたことを特徴とするRNA配列解析装置。

3. RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成

文法とを対応付けて格納する文法格納手段と、

R N A配列を上記生成文法に適用して構文解析木を導出する構文解析手段と、

上記構文解析手段にて導出された上記構文解析木に対して適合度の計算を行う

上記適合度計算手段により計算された上記適合度に基づいて上記R N A配列間の類似度を計算する類似度計算手段と、
を備えたことを特徴とするR N A配列解析装置。

5 6. R N A二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納する文法格納手段と、

R N A配列を上記生成文法に適用して構文解析木を導出する構文解析手段と、
上記構文解析手段にて導出された上記構文解析木に対して適合度の計算を行う適合度計算手段と、

10 上記適合度計算手段により計算された上記適合度のうち予め定めた条件を満たす適合度である上記構文解析木を導出した上記R N A配列を抽出する抽出手段と、
上記構造トポロジーと上記R N A配列とを二次元マトリックスで表示し、上記二次元マトリックスにおいて上記抽出手段にて抽出された上記R N A配列と上記構造トポロジーに対応する格子部分に上記適合度を表示する適合度マトリックスを作成する適合度マトリックス作成手段と、

15 上記適合度マトリックス作成手段にて作成された上記適合度マトリックスについて、上記適合度により上記構造トポロジーをソートし、他のR N A配列について当該ソートされた上記構造トポロジーの順番に対応する上記生成文法により構文解析を行い上記適合度が最大となる上記構文解析木を求め、予め定めた条件を満たす上記適合度を持つ上記構文解析木に対応する上記他のR N A配列を抽出する共通構造抽出手段と、

20 を備えたことを特徴とするR N A配列解析装置。

7. (補正後) R N A二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納する文法格納ステップと、

R N A配列を複数の上記生成文法に適用して構文解析木を導出する構文解析ステップと、

上記構文解析ステップにて導出された上記構文解析木に対して適合度の計算を行う適合度計算ステップと、

上記適合度計算ステップにより計算された上記適合度のうち予め定めた条件を満たす適合度である上記構文解析木を上記適合度が高い順にソートするソートステップと、

上記ソートステップによりソートされた上記構文解析木を上記RNA配列の二次構造の候補として出力する出力ステップと、

を含むことを特徴とするRNA配列解析方法。

10 8. (補正後) RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納する文法格納ステップと、

複数のRNA配列を予め定めた上記生成文法に適用して構文解析木を導出する構文解析ステップと、

上記構文解析ステップにて導出された上記構文解析木に対して適合度の計算を行う適合度計算ステップと、

上記適合度計算ステップにより計算された上記適合度のうち予め定めた条件を満たす適合度である上記構文解析木を導出した上記RNA配列を上記予め定めた上記生成文法が適合する上記構造トポロジーの二次構造を持つRNA配列の候補として出力する出力ステップと、

を含むことを特徴とするRNA配列解析方法。

9. RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納する文法格納ステップと、

RNA配列を上記生成文法に適用して構文解析木を導出する構文解析ステップと、

上記構文解析ステップにて導出された上記構文解析木に対して適合度の計算を行う適合度計算ステップと、

上記適合度計算ステップにより計算された上記適合度のうち予め定めた条件を

行う適合度計算ステップと、

上記適合度計算ステップにより計算された上記適合度に基づいて上記RNA配列間の類似度を計算する類似度計算ステップと、
を含むことを特徴とするRNA配列解析方法。

5

12. RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納する文法格納ステップと、

RNA配列を上記生成文法に適用して構文解析木を導出する構文解析ステップと、

10 上記構文解析ステップにて導出された上記構文解析木に対して適合度の計算を行なう適合度計算ステップと、

上記適合度計算ステップにより計算された上記適合度のうち予め定めた条件を満たす適合度である上記構文解析木を導出した上記RNA配列を抽出する抽出ステップと、

15 上記構造トポロジーと上記RNA配列とを二次元マトリックスで表示し、上記二次元マトリックスにおいて上記抽出ステップにて抽出された上記RNA配列と上記構造トポロジーに対応する格子部分に上記適合度を表示する適合度マトリックスを作成する適合度マトリックス作成ステップと、

20 上記適合度マトリックス作成ステップにて作成された上記適合度マトリックスについて、上記適合度により上記構造トポロジーをソートし、他のRNA配列について当該ソートされた上記構造トポロジーの順番に対応する上記生成文法により構文解析を行い上記適合度が最大となる上記構文解析木を求め、予め定めた条件を満たす上記適合度を持つ上記構文解析木に対応する上記他のRNA配列を抽出する共通構造抽出ステップと、

25 を含むことを特徴とするRNA配列解析方法。

13. (補正後) RNA二次構造の構造トポロジーと、当該構造トポロジーに

適合する生成文法とを対応付けて格納する文法格納ステップと、

R N A配列を複数の上記生成文法に適用して構文解析木を導出する構文解析ステップと、

上記構文解析ステップにて導出された上記構文解析木に対して適合度の計算を
5 行う適合度計算ステップと、

上記適合度計算ステップにより計算された上記適合度のうち予め定めた条件を
満たす適合度である上記構文解析木を上記適合度が高い順にソートするソートス
テップと、

上記ソートステップによりソートされた上記構文解析木を上記R N A配列の二
10 次構造の候補として出力する出力ステップと、

を含むことを特徴とするR N A配列解析方法をコンピュータに実行させるプロ
グラム。

14. (補正後) R N A二次構造の構造トポロジーと、当該構造トポロジーに
15 適合する生成文法とを対応付けて格納する文法格納ステップと、

複数のR N A配列を予め定めた上記生成文法に適用して構文解析木を導出する
構文解析ステップと、

上記構文解析ステップにて導出された上記構文解析木に対して適合度の計算を
行う適合度計算ステップと、

上記適合度計算ステップにより計算された上記適合度のうち予め定めた条件を
満たす適合度である上記構文解析木を導出した上記R N A配列を上記予め定めた
上記生成文法が適合する上記構造トポロジーの二次構造を持つR N A配列の候補
として出力する出力ステップと、

を含むことを特徴とするR N A配列解析方法をコンピュータに実行させるプロ
25 グラム。

15. R N A二次構造の構造トポロジーと、当該構造トポロジーに適合する生

60/1

成文法とを対応付けて格納する文法格納ステップと、

について、上記適合度により上記構造トポロジーをソートし、他のRNA配列について当該ソートされた上記構造トポロジーの順番に対応する上記生成文法により構文解析を行い上記適合度が最大となる上記構文解析木を求め、予め定めた条件を満たす上記適合度を持つ上記構文解析木に対応する上記他のRNA配列を抽出する共通構造抽出ステップと、

を含むことを特徴とするRNA配列解析方法をコンピュータに実行させるプログラム。

19. (補正後) RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納する文法格納ステップと、

RNA配列を複数の上記生成文法に適用して構文解析木を導出する構文解析ステップと、

上記構文解析ステップにて導出された上記構文解析木に対して適合度の計算を行う適合度計算ステップと、

上記適合度計算ステップにより計算された上記適合度のうち予め定めた条件を満たす適合度である上記構文解析木を上記適合度が高い順にソートするソートステップと、

上記ソートステップによりソートされた上記構文解析木を上記RNA配列の二次構造の候補として出力する出力ステップと、

を含むRNA配列解析方法をコンピュータに実行させるプログラムを記録したことを特徴とするコンピュータ読み取り可能な記録媒体。

20. (補正後) RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納する文法格納ステップと、

複数のRNA配列を予め定めた上記生成文法に適用して構文解析木を導出する構文解析ステップと、

上記構文解析ステップにて導出された上記構文解析木に対して適合度の計算を

行う適合度計算ステップと、

上記適合度計算ステップにより計算された上記適合度のうち予め定めた条件を満たす適合度である上記構文解析木を導出した上記RNA配列を上記予め定めた上記生成文法が適合する上記構造トポロジーの二次構造を持つRNA配列の候補として出力する出力ステップと、

を含むRNA配列解析方法をコンピュータに実行させるプログラムを記録したことを特徴とするコンピュータ読み取り可能な記録媒体。

21. RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納する文法格納ステップと、

RNA配列を上記生成文法に適用して構文解析木を導出する構文解析ステップと、

上記構文解析ステップにて導出された上記構文解析木に対して適合度の計算を行う適合度計算ステップと、

上記適合度計算ステップにより計算された上記適合度のうち予め定めた条件を満たす適合度である上記構文解析木を導出した上記RNA配列を抽出する抽出ステップと、

上記構造トポロジーと上記RNA配列とを二次元マトリックスで表示し、上記二次元マトリックスにおいて上記抽出ステップにて抽出された上記RNA配列と上記構造トポロジーに対応する格子部分をマークすることにより、上記RNA配列間で共通に有する構造トポロジーを可視化する共通構造マトリックス作成ステップと、

を含むRNA配列解析方法をコンピュータに実行させるプログラムを記録したことを特徴とするコンピュータ読み取り可能な記録媒体。

25

22. RNA二次構造の構造トポロジーと、当該構造トポロジーに適合する生成文法とを対応付けて格納する文法格納ステップと、

利用者が入力したDNA配列から転写されるRNA配列を作成するRNA配列



PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference PCLA-14604	FOR FURTHER ACTION See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)	
International application No. PCT/JP03/000011	International filing date (day/month/year) 06 January 2003 (06.01.03)	Priority date (day/month/year) 28 December 2001 (28.12.01)
International Patent Classification (IPC) or national classification and IPC G06F 17/30, C12Q 1/68, C12N 15/00, G01N 33/50		
Applicant CELESTAR LEXICO-SCIENCES, INC.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.
2. This REPORT consists of a total of 4 sheets, including this cover sheet.

This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

These annexes consist of a total of 24 sheets.

3. This report contains indications relating to the following items:

- I Basis of the report
- II Priority
- III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- IV Lack of unity of invention
- V Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- VI Certain documents cited
- VII Certain defects in the international application
- VIII Certain observations on the international application

Date of submission of the demand 22 May 2003 (22.05.03)	Date of completion of this report 09 September 2003 (09.09.2003)
Name and mailing address of the IPEA/JP	Authorized officer
Facsimile No.	Telephone No.

I. Basis of the report

1. With regard to the elements of the international application:*

 the international application as originally filed the description:pages 1-4,7,8,11,12,16-50,53, as originally filedpages 5-6/1,9-10/1,13-15/1,51-52/1, filed with the demandpages 5-6/1,9-10/1,13-15/1,51-52/1, filed with the letter of 01 September 2003 (01.09.2003) the claims:pages 3-6,9-12,15-18,21-24, as originally filedpages 1,2,7,8,13,14,19,20, as amended (together with any statement under Article 19pages 1,2,7,8,13,14,19,20, filed with the demandpages 1,2,7,8,13,14,19,20, filed with the letter of 01 September 2003 (01.09.2003) the drawings:pages 1-23, as originally filedpages 1-23, filed with the demandpages 1-23, filed with the letter of 01 September 2003 (01.09.2003) the sequence listing part of the description:pages 1-23, as originally filedpages 1-23, filed with the demandpages 1-23, filed with the letter of 01 September 2003 (01.09.2003)2. With regard to the language, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item. These elements were available or furnished to this Authority in the following language English which is: the language of a translation furnished for the purposes of international search (under Rule 23.1(b)). the language of publication of the international application (under Rule 48.3(b)). the language of the translation furnished for the purposes of international preliminary examination (under Rule 55.2 and/or 55.3).

3. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

 contained in the international application in written form. filed together with the international application in computer readable form. furnished subsequently to this Authority in written form. furnished subsequently to this Authority in computer readable form. The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished. The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.4. The amendments have resulted in the cancellation of: the description, pages 1-23 the claims, Nos. 1-23 the drawings, sheets/fig 1-235. This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).**

* Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rule 70.16 and 70.17).

** Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/JP03/00011

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Claims	1-24	YES
	Claims		NO
Inventive step (IS)	Claims	3, 6, 9, 12, 15, 18, 21, 24	YES
	Claims	1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19, 20, 22, 23	NO
Industrial applicability (IA)	Claims	1-24	YES
	Claims		NO

2. Citations and explanations

The following documents were cited in the international search report.

Document 1: YOSHIZAWA et al., "RNA Niji Kozo Yosoku Shien System no Kaihatsu to Jisso," The Society for Artificial Intelligence Jinko Chino Kisoron Kenkyukai (Dai 32 Kai) Shiryo, March 26, 1998, p. 29-35, particularly, p. 31

Document 2: UEMURA, Y. et al, Grammatically Modeling and Predicting RNA Secondary Structures. Genome Informatics Series, Proceedings Genome Informatics Workshop, December 11, 1995, No. 6, p. 67-76

Based on the description in document 1, the inventions of claims 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19, 20, 22, and 23 lack an inventive step.

Commentary:

Document 1 (p. 31) describes seeking syntactic parsing trees that can describe RNA input sequences and outputting several possible secondary structures for RNA in order starting with those that have the smallest free energy values.

Claims 1, 7, 13, and 19

The description in document 1 also states that it is effective to classify the predicted result by topological structure data (page 34, right column, lines 12 to 14).

Claims 2, 8, 14, and 20

Searching for RNA sequences that can assume given secondary structures is merely a matter to be performed as needed by persons skilled in the art.

Claims 4, 10, 16, and 22

Because cDNA sequences correspond to RNA sequences, persons skilled in the art can easily conceive of converting an input DNA sequence to RNA and predicting the secondary structure as described in document 1.

Claims 5, 11, 17, and 23

Because it can be predicted that RNA molecules will be similar if their secondary structures are similar, equating the goodness of fit of secondary structures to similarity of RNA sequences is merely a matter to be performed as needed by persons skilled in the art.

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/JP03/00011

Supplemental Box

(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of Box V:

The inventions of claims 3, 6, 9, 12, 15, 18, 21, and 24 are novel and involve an inventive step with respect to documents 1 and 2.

Commentary:

Documents 1 and 2 neither describe nor suggest displaying structure topology and RNA sequences in a two-dimensional matrix and marking the lattice members.